

WE CLAIM:

1. A switch for use in a data storage network, comprising:
 - a plurality of ports each comprising a receiving device for receiving data from a link connected to the port and a transmitting device for transmitting data onto another link connected to the port;
 - 5 a plurality of control circuits each associated with one of the ports, wherein each of the control circuits collects data traffic statistics and port state information for the associated port;
 - 10 memory for storing a congestion record for each of the ports; and
 - a congestion analysis module gathering at least a portion of the data traffic statistics and port state information for the ports, performing computations with the gathered port statistics and port state information to detect congestion at the ports, and updating the congestion records for the ports with detected congestion.
2. The switch of claim 1, wherein the module periodically repeats the gathering, the performing, and the updating upon expiration of a sample time period.
3. The switch of claim 2, wherein the congestion records comprise counters for a set of congestion types and the updating of the congestion records comprises incrementing the counters for the ports for which the detected congestion corresponds to one of the congestion types.
4. The switch of claim 3, wherein the congestion types comprise backpressure congestion, resource limited congestion, and over-subscription congestion.
5. The switch of claim 4, wherein the module performs a second gathering of a second portion of the data traffic statistics for ones of the ports for which the detected congestion has the backpressure congestion type of congestion and then processes the second portion of the data traffic statistics to identify a source of backpressure within the switch.

6. The switch of claim 1, wherein the gathered port statistics are selected from the group consisting of TX BB_Credit levels, TX link utilization, RX BB_Credit levels, RX link utilization, link distance, configured RX BB_Credit, queuing latency, internal port transmit busy timeouts, Class 3 frame flush counters/discard frame counters, and destination statistics.

7. The switch of claim 1, wherein the gathered port statistics and port state information include separate sets of data for the receiving device and the transmitting device for the ports and wherein the performing computations comprises detecting congestion for the ports in the receiving device and the transmitting device based on the separate sets of data.

8. The switch of claim 1, wherein the memory further stores a set of congestion threshold values and wherein the performing congestion detection computations with the module comprises determining whether the gathered port statistics and port state information exceed the congestion threshold values.

9. The switch of claim 1, further comprising generating a Congestion Threshold Alert (CTA) indicating one or more congestion statistics to a log or management interface.

10. A method of managing congestion in a data storage fabric having a set of switches with input/output (I/O) ports and links connecting the ports for transferring digital data through the fabric, comprising:

receiving a first set of congestion data from the switches in the fabric, the first set comprising port-specific congestion data for the ports in the switches at a first time;

receiving a second set of congestion data from the switches in the fabric, the second set comprising port-specific congestion data for the ports in the switches at a second time; and

processing the first set and the second set of congestion data to determine a level of congestion at the ports.

11. The method of claim 10, wherein the processing comprises determining a change in the congestion data between the first and the second times.

12. The method of claim 11, wherein the determined change is used to update a set of congestion counters for each of the ports of each of the switches.

13. The method of claim 12, wherein the level of congestion is determined by comparing the congestion counters to threshold levels for a set of congestion types.

14. The method of claim 13, receiving from a user interface at least a portion of the threshold levels and displaying on the user interface at least a portion of the congestion counters.

15. The method of claim 13, wherein the congestion types comprise oversubscription in the receive and transmit directions, backpressure congestion in the receive direction, and resource-limited congestion in the transmit direction.

16. The method of claim 10, further comprising generating a congestion status display for viewing on a user interface comprising a graphical representation of the data storage fabric, the congestion status display including congestion indicators corresponding to the determined levels of congestion at the ports.

17. The method of claim 16, wherein the congestion data comprises detected types of congestion for the ports and the congestion status display includes congestion type indicators.

18. The method of claim 10, wherein the processing includes determining a source of the congestion in the fabric based on the congestion data.

19. A method for managing congestion in a fabric having a plurality of multi-port switches, comprising:

at each switch in the fabric, monitoring bi-directional traffic pattern data for each switch port for indications of congestion and when congestion is indicated for one of the switch ports, updating a congestion record for the congested port based on the monitored traffic pattern data;

operating the switches to transfer at least portions of the congestion records from each of the switches to a network management platform; and

10 at the network management platform, processing the transferred portions of the congestion records to determine a congestion status for the fabric.

20. The method of claim 19, further comprising performing congestion recovery comprising initiating manual intervention procedures or transmitting a congestion alleviation command to one of the switches based on the determined congestion status for the fabric.

21. The method of claim 19, wherein the processing comprises detecting a delta between the transferred portions of the congestion records and a set of previously received congestion records, and further wherein the congestion status comprises a congestion level and a congestion type for congested ones of the ports.

22. The method of claim 21, wherein the processing further includes determining a source of congestion in the fabric based on the types of congestion at the ports.

23. The method of claim 22, wherein the types of congestion comprise backpressure congestion, resource limited congestion, and over-subscription congestion.

24. The method of claim 19, wherein the monitoring at the switches is performed independently in a received direction and in a transmit direction for each of the ports.